Note for inclusion in ATBD Document: The Fifth Advisory Panel Report which

was included as Appendix A of this Response is attached. The Original Appendix

B of this Response is superceded by the Fourth Element of the ATBD Addenda on

the NCAR Model. Therefore, the original Appendix of this response is omitted.

Response to Reviews of the Data Assimilation Office (DAO)

Algorithm Theoretical Basis Document (ATBD).

Richard B. Rood, Principal Investigator

Data Assimilation Office
Goddard Space Flight Center
301-614-6155
rrood@dao.gsfc.nasa.gov

1) Introduction: The Data Assimilation Office (DAO) Algorithm Theoretical Basis Document

(ATBD) presented the comprehensive status and projected developments on all aspects of the

DAO's Goddard Earth Observing System (GEOS) Data Assimilation System (DAS). A

presentation of the ATBD was made to a review board in March 1997. Drs. P. Courtier, R.

Daley, T. Kittel, and S. Lord provided written reviews. This is the response to those written

reviews.

In Dr. Daley's review the DAO activities were divided into seven types: analysis (data

assimilation), atmospheric modeling, acquisition and processing of new data, validation,

distribution of products to community, computing, and management. These seven types will be

used to organize a narrative response of all of the written reviews. By far, the most negative

comments on the scientific-integrity of the DAO algorithms were directed at the atmospheric

modeling. Therefore, more detail will be given on the modeling activities than the other parts.

Of the non-scientific issues, there were consistent comments on issues of computing, data

service, and improved collaboration with other agencies. Some of these issues were addressed in

Chapter 2 of the ATBD. In Chapter 2, as well as explicitly stated in the written reviews, the

successful resolution of these issues lies outside of the DAO. Further, these are issues faced by

all U. S. Earth-science organizations with a mission that requires state-of-the-art supercomputing

and related data systems.   The DAO has been an active and leading member in the community

in the pursuit of solutions to these difficult problems.

2) Analysis (data assimilation): The theoretical foundation of the Physical-space Statistical

Analysis System, PSAS, was found to be solid and possibly "the best implementation of

variational assimilation." Weaknesses were found in some of the practical aspects of turning

solid theoretical work into production algorithms. Since the ATBD review, DAO activity has

been focused on implementation of algorithms, and much of the scientific research in analysis

systems has slowed or stopped (see Appendix A, *Fifth Advisory Panel Report*). There is a

production version of PSAS on the SGI Origin 2000 machines. The research activities have been

focused more strongly on the production system and incremental implementation of successful

research projects into the production system. Production for the Terra system will be with a

PSAS-based system, and backups are being considered in the case that computational challenges

prove untenable.

A parameterized Kalman filter effort has been initiated to accelerate more robust

treatment of statistics for PSAS. The background error correlation model remains only weakly

separable, in the sense that the while the horizontal correlation lengths vary vertically, geopotential and temperature correlation lengths are identical, as are those of the wind and mass-balanced portion of the stream function. Our experience so far with anisotropic, flow-dependent correlations has led us to decide that scientifically meaningful, cost-effective product enhancement will be best achieved by moving away from time- and space- averaged statistics towards a parameterized Kalman filter.

The computational aspects of PSAS, as well as other state-of-the-art analysis systems, are formidable. Scalable software for implementation of analysis algorithms on multiple-processor, distributed-memory computers has proven difficult for all organizations (Seablom[i], personal communication). To determine if the expression of PSAS in physical space provided impossible computational bottlenecks, a workshop, with external reviewers, was held in October 1998 on the numerical aspects of PSAS. The workshop found that the computational aspects of PSAS are not fundamentally different from the computation aspects of the 3-D variational systems used in operational weather centers. With regards to competition with ECMWF, the DAO (and all similar U. S. efforts) are challenged with having to scale to greater than one hundred processors to achieve the same capability the ECMWF achieves with sixteen. Lyster[ii] (1999) argues that the intrinsic complexity of the data assimilation process limits scalability to order 100 processors. Thus, the computational issues will not be reconciled in the immediate future.

3) Atmospheric Modeling: The modeling component of the GEOS DAS has been criticized for a number of years. While the performance of the model was often adequate for the job at hand, this performance was often gained through engineering fixes to algorithm deficiencies. This hurt the scientific credibility of the model and, thus, limited the evolutionary path of the modeling

development. Furthermore, the engineering required to accommodate for the algorithm deficiencies often added many additional computations that reduced computational efficiency and hindered the development of higher resolution models. DAO and Goddard Laboratory for Atmospheres management have long recognized these problems; however, the solution to these problems has not been easy to attain.

When the DAO started, there was a peer-reviewed candidate algorithm for a dynamical core to replace the finite-difference core of the GEOS-1 model. This was the Bate's et al. (1993) semi-Lagrangian core[iii]. A complete model was developed; however, this candidate model was not chosen to replace the existing GEOS-1 core for scientific, computational, and programmatic reasons. Thus, this research effort was ended and did not lead to a state-of-the-art replacement model for the DAO.

The performance and completeness of the physics parameterizations in the GEOS DAS have also been the subject of criticism. Both internal (DAO and GSFC) and external (NCEP) initiatives were started to address the shortcomings in physics. Some of these efforts have been successful, but in general it has proven difficult to provide a complete suite of physical parameterizations that are broadly accepted by the community as up to standards. As stated in the ATBD, we concluded that a new process was necessary if a product-oriented effort such as the DAO was to be dependent upon external scientific development for mission critical components of the data assimilation system. Furthermore, we concluded that the existing internal efforts in NASA could not provide the state-of-the-art components necessary for the data assimilation effort and that external partnerships were necessary.

In the ATBD review, DAO management was stated to demonstrate "intransigence" on the model issue. Based on the above experience and observation of and discussions with other

groups who have imported models, DAO management (*i.e.* R. Rood) felt that the recommendation to simply import a model assured that the DAO would maintain a second-class modeling effort. Importing and maintaining a model from another organization, without the infrastructure in place, without the commitment of the other organization to maintain a state-of-the-art model for DAO needs, assures that the DAO will always be maintaining a model that is a step behind. Importation does not provide what is really needed – robust scientific development of model components focused on the DAO application.

DAO management (*i.e.* R. Rood) maintained that if modeling (or any mission-critical component of the DAS) is to rely on outside resources, then a partnership is needed. The ATBD reviews increased the urgency for the DAO to address the modeling issue, requiring that a radical new approach be initiated to assure that a new state-of-the-art model is available for the GEOS DAS. With this added urgency and motivation from Division-level management at both NCAR (M. Blackmon) and GSFC (F. Einaudi), D. Williamson and R. Rood set out to develop a successful collaboration to unify model development between the DAO and the Climate and Global Dynamics Division at NCAR. This effort based on a next-generation dynamical core provided by the DAO and Community Climate Model (CCM) physics package provided by NCAR is described in Appendix B. This includes a discussion of the elements of a successful collaboration.

A working model has been built, and initial assessment of the model performance at both DAO and NCAR shows surprising robustness. DAO has made a commitment to this model as its next assimilation model (PM-1 and beyond). Serious discussions are underway that the next generation of the Community Climate Model (CCM4) will be the joint model. As detailed in Appendix B, the joint model addresses some of the fundamental problems present in CCM3 and

its candidate replacement. Since the writing of Appendix B, initial forecast experiments have

been successfully completed with the joint model (Figure 1). Animations of the breakup of the

stratospheric vortex and model-generated hurricanes can be found on the DAO web site[iv].

4) Acquisition and Processing of New Data: The DAO views the use of new types of

observations as a centerpiece of its mission. The design and development of PSAS is

specifically directed at the best use of new data, including non-state variables.   In addition,

traditional observations from both satellite and 'conventional' sources must be effectively and

thoroughly utilized. The assessment of the impact of new data types needs to be made on the

background of the traditional observations in order for the results to be accepted by the scientific

community. The scientific, computational, and management challenges of use of new data types

is as difficult part of the data assimilation process as any other activity.

Since the ATBD review, there has been significant progress in assessing the impact of

new types of data. The work with observations from the NASA scatterometer received wide

recognition and accelerated the use of scatterometry data by the NOAA Marine Prediction

Center. In numerical weather prediction applications the scatterometry data provided a full day

improvement in forecast skill in the Southern Hemisphere, the largest forecast improvement ever

realized by a single space-based platform. Assimilation studies with TRMM observations have

already shown improvement in both prognostic capability and determination of climate

parameters[v]. The DAO has field-leading efforts in ozone assimilation as well as other types of

constituent observations. The constituent work includes collaborations with both Duke

University and the NCAR for assimilation of MOPITT data. In continuing to improve of the

treatment of the operational TOVS data, the DAO has developed algorithms for the SSU

component of TOVS. The SSU algorithm has been requested by ECMWF for its reanalysis. There are also active efforts in the assimilation of several other sensors including SSM/I, the improvement in the treatment of traditional data sources, and collaboration with Meteo France on the assimilation of GPS observations.

The DAO has also vastly improved its capability to monitor and quality control the input data stream. Monitoring is the first step in incorporating new types of observations. It is necessary to characterize the statistics of the observations as well as the ability of the model to represent the observation type. The monitoring system as applied to the addition of the SSM/I surface wind observations is exemplified on the DAO web site[vi]. Improvement in quality control algorithms has benefitted from collaboration with NCEP.

Since the ATBD review, the DAO has performed impact studies that evaluate the geophysical information in new observations, proven the ability of the assimilation system to use that information, and accelerated the use of data by other agencies. This is an important part of the DAO mission. Also, since the ATBD review, several problems have been revealed that increase the risk that the DAO will not be able to meet expectations on the use of new data. In general, these risks are not due to scientific shortcomings, but are due to computational and administrative insufficiencies. The solution to these problems is not always in the control for the DAO to address.

The administrative problems are of two types. The first type involves data agreements with internal NASA projects and external data providers to assure reliable, timely data streams. The Memoranda of Understanding and the Interface Control Documents, undertaken by both the DAO and ESDIS, are often fragile and incomplete. Several times interruption or unexpected

changes in the input data stream have caused failure of the DAS and slowed the delivery of expected products.

The second type of administrative problem involves prioritization of the incorporation of new data types. For example, there is significant funding for the assimilation of carbon monoxide; however, there is no specific funding for assimilation of SSM/I observations. The SSM/I observations are deemed more critical by the bulk of DAO customers and are fast becoming a part of the standard observation set. In the case where curiosity-driven or program-driven initiatives determine the primary path of which new data sets to assimilate, the development of the system as a whole tends to be fractured. A more robust mechanism to assure the assimilation of the fundamental observations is needed, and R. Rood has brought this issue to the attention of the EOS Atmosphere Panel. If a better way to assure that the infrastructure datasets are properly accounted for can be developed, then the proposal-driven activities can be more appropriately directed towards discovery research with new instruments.

The computational aspects of the use of new types of observations bring to the forefront the complexity issues of data assimilation. Indeed, the current data sets are large and tax the computational resources available for storing and moving the data sets. The projected future data sets are much larger. Further, the data sets are heterogeneous; therefore, commercial technology developed for transactional databases have only limited applications. The size of the datasets motivates scientific research into extracting the maximum amount of information from the minimum volume of data prior to assimilation. The diversity of datasets increases demands on central-processor requirements as more and more datasets are intercompared in the quality control portion of the DAS. Since the DAO is still reliant on shared-memory computers, new data can tax the memory limits and have, in fact, sometimes crashed the DAS. The increasing

complexity and volume of input data sets demand reliable, balanced interactions in complex computer systems. The ability to move beyond impact studies to routine assimilation of new data types requires more effectively addressing the data system issues associated with both NASA and non-NASA data management.

With the recognition that many of the challenges of use of satellite data are common to all assimilation efforts, from 21-23 April 1998 DAO hosted a "Workshop on Assimilation of Satellite Data." R. Errico of NCAR chaired the meeting with G. Ohring (NESDIS). The workshop included organizers from all of the major centers in the U. S. The workshop report[vii] recognized many general problems in the assimilation of satellite data and stated that progress had in general been "slow." "Its (the assimilation of data) acceleration requires augmentation of several resources: computational, personnel, infrastructure, and basic research." The community clearly recognizes the need for the agencies to work together more effectively. This is one research area where collaboration with NCEP should be most useful, and initial contacts have been made.

5) Validation: Since the ATBD review the DAO has formalized its validation process using the validation plan presented in Chapter 8 of the ATBD as a guideline. Historically, validation was driven by discovery-based research as the original funds for validation came through individual scientist's RTOP proposals. Currently the DAO has four categories of validation based on their timeliness: near-real time, short-lag, medium-lag, and long-term. The near-real time validation is largely automated with tests for unrealistically high and low values of basic parameters as well as flagging changes in characteristics of input data acceptance[viii]. The short-lag validation is anchored by the DAO On-Line Monitoring System (DOLMS). DOLMS is a graphical display of

basic geophysical and statistical parameters, often with side-by-side displays of time-averaged quantities or comparable products from other sources. DOLMS is web-based[ix] and is generally available to all users at Goddard. DOLMS is made available to external customers on a request basis.

The medium-lag validation is "system validation" and is focused on specific candidate systems compared with existing baselines. Validation committees are formed to evaluate geophysical parameters. These parameters are determined either by specific shortcomings in the baseline system or by system enhancements aimed at the inclusion of new parameters (*e.g.* adding explicit cloud liquid water). Both the identification of shortcomings and scientific development paths are guided by internal and external customer feedback. It is also the responsibility of the system validation team to assure and document that the integrity of the baseline product is maintained.

Weather forecasting is becoming a larger part of DAO activities and is a key part of the system validation effort. Historically, within the DAO, the role of weather forecasting has been controversial because of potential mission conflict with other agencies. However, NASA is now trying to improve links with the forecasting community and facilitate the use of new data types by operational agencies responsible for weather forecasting.

Furthermore, the ability of the assimilating model to perform credible weather forecasts is a standard demanded by many users of assimilated data sets. Forecasting provides one timely measure of data impact and is a proven integrated monitor of the integrity of the input data stream. Further as the modeling field has evolved, weather forecasting and climate simulation needs to be increasingly integrated to assure physical consistency. Of particular importance to the DAO's forecasting activities are meteorological realism, for instance, the ability to represent

fronts (Figure 2), and the behavior of the physics parameterizations in the presence of observational forcing.

The long-term "scientific validation" focuses on complex Earth-science problems and remains, basically, curiosity-driven. The long-term validation assures that applications that stress the internal consistency of assimilated data sets are intimately linked to the development process. The internal efforts feed back more quickly and effectively than external efforts, whose results are usually gathered through *ad hoc* processes or user workshops[x]. It is critical that the DAO have internal scientific validation capabilities to maintain a vital intellectual environment and to assure that the GEOS data assimilation system development with the guidance of real problems.

In the ATBD reviews issues of the land surface were particularly noted. The DAO, jointly with the Hydrological Sciences Branch (Code 974), formed a standing team to work on land-surface problems. An off-line assimilation system was built to provide flexible and timely research. The Off-line Land-surface GEOS Assimilation (OLGA)[xi] effort has focused on sensitivities of the land-surface model and improvement of the meteorological parameters that serve as the boundary conditions for the land-surface model and assimilation system. Improvement of the atmospheric boundary conditions for land-surface applications is a first priority problem. Simultaneous with the OLGA effort, the Koster and Suarez[xii] (1996) land-surface model has been incorporated into the core assimilation system. This will be part of the DAS used to support the Terra launch. The interactive land-surface model addresses many of the problems with surface temperature biases (Figure 3). The May 1998 Fifth Advisory Panel Report (Appendix A) notes the substantial improvements in the land-surface capabilities of the DAO.

Of particular note for future efforts are statistical validation methods. These methods become possible with more advanced data assimilation systems that rigorously treat the error characteristics. Statistical validation approaches test the very assumptions made about the error sources of the different instrument types as well as how these error characteristics propagate through the assimilation system.

6) Distribution of Products to Community: The DAO has been significantly impacted by the well-documented problems with the development and deployment of EOSDIS. The DAO has made provisions to meet the needs of the instrument teams and campaign teams that we support with our own resources. Distribution to the broader community remains *ad hoc*, with some collaborative efforts either being proposed or implemented by other groups. Data service is critical for data usage and feedback from customers. This service needs to be improved. The scope of the problem is broader than the DAO can address alone.

7) Computing: Computing has been explicitly identified as the highest risk to DAO success since the 1994 plan. This remains the case in 1999. As pointed out in the ATBD reviews and several of the DAO Advisory Panel Reports, the computer problem is not unique to the DAO. It is, in fact, common to all organizations in the United States that require the use of supercomputers for time-critical product delivery. Earth-scientists in the United States are at a distinct disadvantage relative to our foreign counterparts because the U. S. Agencies are restricted in their acquisition of computers. For a given level of performance, the machines available to scientists outside of U. S. are much easier to use, with the machines available in the U. S. requiring an enormous software development. Further, it is not clear whether such a

software development will yield successful scalable code (see Section 2). Therefore, the gap between product-oriented Earth-science activities in the United States and the rest of the world is increasing.

The strategy that the DAO has undertaken with NASA/Ames Research Center came from necessity. DAO production for Terra will be at Ames, and, all things considered, the partnership is successful. Unequivocally, the ability of the DAO to deliver products for Terra comes from the partnership with NASA/Ames. However, the computing systems are complex and balky, and computing across the country is not seamless.

As pointed out in the ATBD reviews, corporate commitment to high performance computing is unsure. The DAO has modified and restricted its science development to meet mission support requirements with available hardware and software and within budget. In order to provide the computing resources to maintain the standards being set by foreign efforts, a significant increase in computing resources is required. This includes a directed effort for applications software development as well as platform middleware that used to be provided by the vendors. This sort of investment requires a strategic decision at the programmatic level, and significantly more productive intercenter and interagency collaborations.

There are no easily prescribed solutions to the computing problem. R. Rood has web-published "Elements of a Supercomputing Strategy"[xiii], and currently is on special assignment to NASA HQ and several interagency committees to develop a strategy to address the software and hardware issues confronting Earth Science. The above document is being updated to identify specific investment needs as well as clarification of the relationship between different applications of global models (*i.e.* data assimilation, weather forecasting, seasonal prediction, climate modeling, etc.).

8) Management: First, R. Rood would like to thank Dr. Daley for his compliment of DAO management. R. Rood remains the Principal Investigator of the effort, but has turned over the day-to-day management of the DAO to Dr. R. Atlas. R. Rood remains intimately involved with the DAO, working largely on developing an external environment more conducive to DAO success. This includes development of an integrated modeling and computing strategy with the identification of the infrastructure investment needed to make a high-end assimilation effort sustainable. He is totally committed to the notion that success of the DAO depends on interagency partnerships to address computing, modeling, data use, and assimilation problems. Such interagency partnerships are a new way of doing business, and successful business models are difficult to define. R. Rood is beginning a one-year special assignment to NASA HQ to specifically address these issues.

Much attention was paid in R. Rood's final year of being the Head of the DAO to the development of a viable software group, as well as a more structured approach to science management. Many of these changes were initiated with the advice of external experts in organizational management. Drs. Atlas, Hou, Bloom, and others have continued to try and develop this culture. They should be contacted directly for more information.

Figure 1. Results from initial forecasts experiments with joint DAO – NCAR model. This figure shows 500 mb geopotential height anomaly correlations from initial forecasts with the joint DAO-NCAR model. These studies were carried out to assure that nothing in the model formulation yielded the model unsuitable for forecast applications. Results from three forecast models are presented, GEOS-2, the current finite-difference model used for DAO production, the NCEP model as used in the NCEP reanalysis, and the joint model using the Lin-Rood control-volume dynamical core and the NCAR physics. Since the initial forecasts were focused on model viability, significance should not be attributed to the forecast skill. However, in all experiments the joint model shows promise in forecast applications.

Figure 2. Representation of surface front in joint DAO – NCAR model. This figure shows the representation of a simulated (not forecast) surface front from the joint DAO-NCAR model. One advantage of the control-volume dynamical core is the ability to increase the realism of model simulations. This representation of a front has the characteristics of fronts as observed in nature, including the discontinuity in wind direction across the frontal boundary. In the spectral models used at operational centers, fronts are generally represented only as troughs in the pressure fields. With the accurate simulation of fronts, better use of high resolution observations, such as from the NASA scatterometers, becomes possible. Such realism also motivates additional research into flow dependent correlation models, as the influence of observations on either side of the front is obviously anisotropic.

Figure 3. Impact of the land-surface model on representation of clear-sky outgoing longwave radiation. This figure follows from validation exercises in cooperation with the CERES/TRMM team. Shown is outgoing longwave radiation (OLR) over two types of surface. The black line is the observation from CERES. The red line is the assimilation product using the GEOS-2 DAS without an interactive land surface model, the blue line includes the land-surface model. The top two figures show the OLR and the difference from the estimates from the assimilation from CERES over crop land. The bottom two panels are over desert. In both cases the bias is reduced.

FIFTH REPORT OF THE

DATA ASSIMILATION OFFICE ADVISORY PANEL

TO THE

LABORATORY FOR ATMOSPHERES

NASA GODDARD SPACE FLIGHT CENTER

R. Daley, Chairman
Naval Research Laboratory (Monterey)

J. Anderson
Geophysical Fluid Dynamics Laboratory

A. Bennett
Oregon State University

P. Courtier
LODYC CNRS (France)

R. Dickinson
University of Arizona

D. Jacob
Harvard University

D. Johnson
University of Wisconsin

J. O'Brien
Florida State University

A. O'Neill
Reading University (UK)

MAY 11-13, 1998

INTRODUCTION

The Data Assimilation Office (DAO), after several years of substantial human and monetary investment, has reached the stage where it can truly begin to fulfill its mission--the production of research-quality assimilated data sets for the earth sciences from the heterogeneous array of in situ and remote EOS instruments. For the DAO, this is a complete change from its pre-1992 focus on the production of purely meteorological data sets for short range weather forecasting.

The early years of DAO (1992-1995) were devoted to consolidating several groups from within the Goddard Laboratory for Atmospheres (GLA) and re-focussing their efforts towards the DAO mission. A new structure was created based on modeling, data assimilation and validation and the DAO culture has been gradually re-oriented from curiosity-driven to mission-oriented science. Outstanding new scientists were hired during this period and DAO became an extremely innovative environment in which to work.

AM-1 launch has served for 5 years to focus DAO activities, although it has been delayed until late 1998. In order to provide a credible data assimilation system for AM-1 launch, DAO has been recently split into a research component and an operational component, with a strong emphasis on science in the former and engineering in the latter. The operational system (GEOS-2) is now up and running.

The Advisory Panel was formed in 1992 to help guide the DAO in its early stages. Reports produced by the panel in December 1992, December 1993, January 1995 and June 1996 made a number of recommendations on organizational, personnel, computer, and technical matters. The panel's recommendations have always been given serious consideration by DAO and many of them have been successfully implemented over the past few years.

The Advisory Panel visited DAO for the fifth time on May 11-13, 1998. There were two purposes for this visit. First, to be assured that DAO was operationally ready for the AM-1 launch and secondly, to explore how DAO will evolve once it becomes fully operational. All members of the panel were present. The first day was devoted to presentations by DAO personnel on the prepara-ions of the GEOS-2 system for AM-1 launch. The remaining time was devoted to scientific presentations by DAO personnel, preparation of major recommendations, and a briefing of DAO and NASA Goddard management. A draft of this report was prepared and circulated to members of the Advisory Panel for comment. The report will discuss the progress of the DAO since the last Advisory Panel meeting, the operational system for AM-1 launch, and present and future activities. The report will conclude with a series of specific recommendations.

DEVELOPMENTS SINCE LAST PANEL REPORT

We first reviewed the progress made by DAO since the last visit of the panel in June 1996. This last visit was almost 2 years earlier and there had been considerable activity in the interim.

Major Organizational Changes

There is a considerable difference between the organization of operational centers such as ECMWF and NCEP and the quasi-operational configuration that DAO had had for some time. It is true that DAO usually has (except when supporting experiments) the luxury of working at some time lag. This removes the complication of a round the clock operation, but many of the other stringent requirements of an operational system remain. These operational requirements dictate certain universal organizational principles, which must be followed if the mission is to be completed successfully.

By far the most important of these principles is that a data assimilation center is a hybrid with both research and operational (production) functions. Both the research and operational functions are equally important, but they must be effectively separated. In other centers this separation is achieved in different ways, some more successful than others. In the research function, a strict scientific modus vivendi should prevail, while in the operational function a more engineering culture is appropriate.

These two functions should be governed by appropriate line management, with the lines meeting just below the level of the Head of DAO. It goes without saying, that leadership on both sides is important at all levels. Interaction, both formal and informal, between the two functions would be encouraged at all levels.

It is difficult to switch from an entirely scientific culture to this dual arrangement and it was anticipated that the switch, when it came, would not be painless. At DAO the crisis came in the fall of 1997, and, hopefully, the worst is now over.

Computation

DAO has always suffered and will continue to suffer from computational problems which are not of its own making and are to some extent beyond its control. Powerful vector machines such as those used at ECMWF are unattainable in the U.S., and the only alternatives must be chosen from an ever-changing menu of unappetizing distributed memory machines marketed by uncommitted vendors. Of course, DAO shares this particular problem with other U.S. groups, but there seems to be less understanding and appreciation within NASA for the computational function.

DAO has tried to solve this problem in a very imaginative fashion, the cross-continental link to AMES. As noted in an earlier report, this was a risky strategy which had never been successfully attempted for a full-blown data assimilation system. All indications are that the

system is working reasonably well, and will continue to work as long as Ames remains committed.

The re-organization of DAO into operational and research components, together with other factors, has had an impact on software development, particularly on issues such as coding standards, MPI conversion and configuration control.

Production

The gradual emergence of the operational or production function of DAO has resulted in new capability and some operational successes. Clearly, mission support from DAO for stratospheric missions is at an increasingly high level. Support for CERES/TRMM, while in its early stages seems to have been successful and will provide useful lessons for AM-1.

Research from the validation, data assimilation and modeling groups is starting to transition into a production monitoring function. In particular, input (raw data) monitoring and output (characteristics of products) has started in earnest. While DAO has been unable to secure a dedicated briefing room (difficult with personnel so widely spread), they have partially got around this problem with a web-based virtual briefing room called (DOLMS-DAO on-line monitoring system).

Research

The following important scientific advances have been made:

Stratosphere

Stratospheric data assimilation is becoming increasingly important and DAO is at last becoming a world leader in this field. Stratospheric modeling, theory and observation have always been strong within Code 910, and now there is credible data assimilation capacity to match it. The extension of GEOS-2 and developments in stratospheric covariance modeling have resulted in a much more credible stratosphere/mesosphere products. Initial experiments with ozone assimilation look very promising. The validation effort is making increasing use of stratospheric diagnostics.

Land surface modeling and assimilation

Land surface aspects of the GEOS-2 system have been substantially improved and initial efforts are being made to assimilate land surface quantities. A number of talented scientists have focussed on this effort, a version of MOSAIC has been implemented in the GEOS-2 model, and extensive validation has been carried out. Overall, the present direction should lead to major future improvements.

New instruments

Assimilation of observations from new instruments is an important part of the DAO mandate. Considerable progress has been made in the direct assimilation of radiances from TOVS, in preparation for the launch of the hyperspectral sounder AIRS. The assimilation of scatterometer data from NSCAT has also been successful. Experiments on the assimilation of SSM/I precipitable water and rainfall are also underway.

New dynamic core

The Lin/Rood dynamic core development has progressed very rapidly. Recent cooperative experiments with NCAR look very promising.

Re-analysis

The re-analysis effort based on the GEOS-1 system provided many useful lessons. It demonstrated how the validation effort could be integrated with modeling and assimilation to improve modeling and/or assimilation components in addition to learning about the workings of the climate system.

ANALYSIS OF THE PRESENT SITUATION

There are basically two issues that are important to DAO at the present time--operational readiness for the AM-1 launch and a credible way forward into the next century. Both issues contain operational and scientific components.

Organization

As noted earlier, DAO has begun the metamorphosis into a truly operational organization, with the cleavage between operations and research gradually becoming a reality. This is a very difficult process and there are both external and internal obstacles to be overcome.

Externally, this type of mission-oriented combined operational/research organizational structure, somewhat paradoxically, does not seem to fit easily into the NASA system and culture. As noted earlier, both the scientific and engineering culture have to thrive and co-exist within the DAO, if it is to succeed in its mission. This problem needs more attention from upper NASA management.

This re-orientation of DAO also suffers from internal resistance. This is particularly evident among scientists, who do not seem to really appreciate the importance and difficulty involved in turning the DAO boutique assimilation system into a robust, efficient production system. There is an inability to hand off developmental software to software people when the time has come for implementation. This latter problem is due partly to the attitudes of the scientists themselves and partially because of the lack of appropriate management mechanisms.

There are important key positions in the line management on both the scientific and operational sides, which are either unfilled or filled with inexperienced managers. DAO is still forced to rely far too heavily on its Head (R. Rood). The proposed organizational chart seems very vague.

Medium and longer range planning, which was strong 2 years ago, seems to have languished, although shorter range planning is okay.

Computing

Computing will always be of concern as long as DAO does not have access to a really fast machine. DAO seems committed to SGI Origin 2000 machines, but the throughput from these machines still seems inadequate for the task at hand. The ability to only do 3-5 days of assimilation per day is just not adequate and gives insufficient resources for research; nor does it provide an adequate buffer in case of catastrophic failure. Support by the vendor is disappointing.

While the AMES connection seems to work better than might have been expected, it still seems rather fragile and subject to changes in NASA policy. The bandwidth for transmitting the fields back and forth between AMES and Goddard seems adequate.

There is the possibility of cooperation with the ASCI project of DOE. This may or may not lead to useful assistance with these computing difficulties.

Software development has suffered from the delay in uncoupling the research and operational functions. In particular, MPI conversion has really lagged, which is another reason why the GEOS-2 system runs too slowly. Other aspects of the software such as standards and configuration management are slowly improving, but have some distance to go yet.

Modeling

While there has been progress in modeling, particularly with respect to land surface parameterization and the stratosphere, the old, nagging problem remains. That is, the dynamics core of GEOS-2, which is slow, based on old technology and with a pole problem that just won't go away.

There may be a long term solution to this problem, through interaction with NCAR. The Lin/Rood model developed at DAO has a number of features which would be attractive to both NCAR and DAO. It employs modern numerics, has no pole problem, takes a long (semi-Lagrangian) time step, and is easy to run on distributed memory machines. It has been successfully integrated at NCAR for long time periods using NCAR physical parameterizations. There may be assimilation problems with this model, however, due to the vanishing layers. A joint modeling project with NCAR could offer considerable benefit to DAO, but should be approached with some caution.

Stratosphere

There has indeed been a lot of progress here as noted previously. However, other resources in Code 910 have still not been tapped. While there has been some success with constituent assimilation of some UARS data (CLAES and MLS) using the Kalman filter, this has not yet translated into a more practical stratospheric assimilation of UARS data.

Much of the UARS data set is obtained from limb sounders, which are (probably) more difficult to assimilate than the nadir sounder data upon which most of the effort in the data assimilation community (including DAO) has been expended.

Assimilation System (PSAS)

It has taken a considerable time to tune this system for operational use, although the tuning seems to be essentially completed. Insertion of quality control is still not finished. The MPI conversion lags, even though an early version was done at JPL more than three years ago. This has meant that the code is running much more slowly than it should.

This particular project has suffered more than others from the lack of a clear demarcation between research and operations. Scheduling and priorities have been chaotic and scientists seem to have often been working at cross-purposes. Operational priorities have too often been subordinated to scientific priorities.

Over the past 5 years, the DAO data assimilation group has been (with its counterpart at ECMWF) the strongest data assimilation group anywhere. Enormous contributions have been made by its members to many aspects of data assimilation. The emphasis on getting PSAS into operational use (which has, of course, been strongly endorsed by this panel in the past) has had one serious adverse effect. That is, the theoretical, investigative aspect of the work has been very downplayed and the work being done is much less exciting than before. Perhaps this is unavoidable, but there is real danger that DAO may be mortgaging its future.

RECOMMENDATIONS

(1) Organization - DAO's reorganization into separate production and research functions should be pursued vigorously until it is completed. It is absolutely necessary that this initiative succeed and it will only succeed if DAO management is strongly supported by both higher NASA management and all DAO personnel.

(2) Personnel - the success of DAO is far too dependent on one person--Ricky Rood. The line management structure must be beefed up with stronger line managers. This is particularly true with the new structure mentioned in (1). Separate science and operational managers must be developed or brought in and empowered.

There are still a few holes that should be filled on the operational side (software, in particular).

(3) Planning - there is no clear idea where DAO is going over the 5 and 10 year range. Medium and long range planning should be revived to consider such questions as the future after AM-1 launch and the continuity of the organization after Ricky Rood steps down.

(4) Operations - should continue along its present track. Leadership problems (as noted in 2 and 3) must be addressed. The virtual briefing room is a brave attempt to circumvent the problems of non-collocated personnel, but there is no substitute for the real thing.

(5) Computing (hardware) - DAO should continue with its SGI Origin based production computing at AMES. This is despite the fact that the SGI Origin has not come up to expectations, nor has the support from SGI been adequate. DAO should NOT be tempted to use the GSFC T3E for production, if it should happen to come available.

(6) Computing (software) - the DAO GEOS-2 system is slow and the throughput is completely inadequate for a true production system. The main obstacle to adequate throughput from the production system is the lagging MPI conversion. This conversion must be pushed strongly or the DAO system will never be able to ingest enough observations, run at sufficiently high resolution or take advantage of computing or scientific opportunities as they arise. This is particularly important, because of the obstacles on the hardware side. Software standards must be improved and software specialists and managers given more respect from the research scientists. Appropriate management procedures must be put in place to ensure a smooth transition of software from research to operations.

(7) DOE initiative - there is the possibility of obtaining some aid for DAO's software and hardware problems through the DOE climate initiative. This possibility should be investigated.

(8) NCAR - the initiative with NCAR should be cautiously pursued. It clearly has potential benefits for both DAO and NCAR. This initiative will not succeed if the interaction is only lower-level and scientific. It requires backing (perhaps contractual) from the upper management of both NASA and NCAR/NSF. There are many questions that must be considered, such as whether this combined modeling approach would be acceptable to the NCAR CSM community, or how compatible this NCAR/DAO modeling initiative would be with present attempts to rationalize the chaotic U.S. modeling situation.

(9) Modeling - DAO cannot continue indefinitely with the present GEOS model. Concrete steps must be taken to secure a new model. The Lin/Rood model, together with other new dynamic frameworks should be assessed, possibly with the help of NCAR (see 8). A candidate model should be chosen after an appropriate validation and review process and (hopefully) supported by both institutions.

The resulting model would have to be flexible enough to serve both institutions. This means a comprehensive physics parameterization library from which could be selected appropriate algorithms for the high vertical and horizontal resolution requirements of DAO or for the lower resolution required for NCAR climate studies. The horizontal resolution has to be better than 50

km, and it should have a good hydrological cycle, prognostic clouds and liquid water, and a good stratospheric/mesospheric simulation.

(10) Stratosphere and chemistry - stratospheric data assimilation is an important new field, in which DAO could and should become the acknowledged leader. This will require a sustained commitment, both within DAO and in other sections of Code 910. Greater effort should be put on the assimilation of UARS data, in fact, a re-analysis project based around UARS data would be extremely useful. Greater emphasis (both practical and theoretical) should be placed on the assimilation of limb sounder data, in general.

More work is needed on the assimilation of chemical constituent data. After ozone has been successfully assimilated, the next candidate tracer should be N_2O. Chemistry tends to be oversimplified in present applications and off-line work should be done with more sophisticated chemistry (ozone depletion and heterogeneous chemistry).

(11) Hydrological cycle - support of the TRMM mission should pay off for both DAO and TRMM. Comparisons of cloud cover from model and assimilation system with satellite observations on short time scales would be useful.

(12) Land surface processes - two meter temperatures should be collected and compared with model and assimilation products. Eventually this information should be assimilated as well. It is also timely to start working toward assimilation of PM-platform microwave observations.

(13) Sea surface - the techniques developed by Bob Atlas' group for the assimilation of scatterometer winds should be incorporated in the main GEOS data assimilation suite. In the meantime, a two dimensional version of PSAS should be run offline to assimilate scatterometer, SSMI, and other surface data to produce high quality surface wind stress fields over the oceans.

(14) The GEOS-1 system - the quality is so far below present day standards that it should not be run any more.

(15) Re-analysis with GEOS-2 - although this would be desirable, it should not be attempted until there are adequate computer resources.

(16) Review panel - it is time to think of reconstituting this review panel.

---

[i] Summer-Fall of 1998 M. Seablom of the Earth and Space Data Computing Division at NASA/Goddard surveyed NCEP, ECMWF, and the Canadian Meteorological Center to determine the computational aspects of analysis systems for global assimilation. At that time ECMWF was still scaling to only 16 processors and running their analysis at lower resolution than the model. The other centers had not developed schemes for distributed-memory parallel architectures. R. Rood regularly checks progress at both ECMWF (D. Dent) and NRL Monterey (T. Hogan). At these two institutions scalable versions of the analysis systems are far behind scalable versions of the model.

[ii] http://dao.gsfc.nasa.gov/DAO_people/lys/complexity

(Lyster, 1999, submitted *J. Scientific Programming*) In this paper an Amdahl's Law argument is put forth. Basically, due to the complexity and intermittency of the assimilation process, it will be difficult to remove sequential steps from the system. Therefore, the scalability is intrinsically limited. Reduction of the sequential steps of the system to 1 % would yield a scalability limit of 100.

[iii] Bates, J. R., S. Moorthi, R. W. Higgins, 1993, A global multilevel atmospheric model using a vector semi-Lagrangian finite-difference scheme 1. Adiabatic formulation, *Mon. Wea. Rev.*, 121, 244-263.

Moorthi, S., R. W. Higgins, J. R. Bates, 1995, A global multilevel atmospheric model using a vector semi-Lagrangian finite-difference scheme 1. Adiabatic formulation, *Mon. Wea. Rev.*, 123, 1523-1541.

[iv] http://dao.gsfc.nasa.gov/DAO_people/lin

[v] http://dao.gsfc.nasa.gov/ScienceEfforts/Highlights/TRMM

[vi] http://dao.gsfc.nasa.gov/monitoring/eval/ssmi01_w98g/qc/

[vii] Errico, R. M., 1999, Workshop on Assimilation of Satellite Data, *Bull. Amer. Met. Soc.*, 80, 463-471.

[viii] Click on Contents, then Quality Assurance of Data Sets (QuADS) of ftp://niteroi/www/on9706/on9706.html

[ix] http://dao.gsfc.nasa.gov/monitoring

[x] First WCRP International Conference on Reanalyses, 27-31 October 1997; Workshop on GEOS-1 Five-Year Assimilation, 6-8 March 1995

[xi] http://land.gsfc.nasa.gov/OLGA/OLGA_Web.htm
http://dao.gsfc.nasa.gov/DAO_people/houser/l1/papers/OLGA.html
http://dao.gsfc.nasa.gov/DAO_people/mikeb/page11.html

[xii] NASA Technical Memorandum 104606, Vol. 9, (Energy and Water Balance Calculations in the Mosaic LSM, Randal D. Koster and Max J. Suarez, 1996)

[xiii] http://dao.gsfc.nasa.gov/Publications/supercomputing.html